



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 147098**

**TO: Nita M Minnifield**  
**Location: rem/3c01/3c18**  
**Art Unit: 1645**  
**Thursday, March 10, 2005**  
**Case Serial Number: 10/613228**

**From: Barb O'Bryen**  
**Location: Biotech-Chem Library**  
**Remsen 1A69**  
**Phone: 571-272-2518**

**barbara.obryen@uspto.gov**

### **Search Notes**

# RUSH

*Reviewed  
3/17/05  
mm*

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 16:53:39 ; Search time 156 Seconds  
(without alignments)  
230.757 Million cell updates/sec

Title: US-10-613-228A-1

Perfect score: 22

Sequence: 1 tcgtcgtttttcgtcggtttt 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_NA.\*

- 1: /cgn2\_6/ptodata/1/ina/5A-COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5B-COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A-COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B-COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/PCTUS-COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	17.2	78.2	133	4	US-09-313-294A-108
2	17.2	78.2	513	4	US-09-107-532A-292
3	16.8	76.4	36016	4	US-09-949-016-14223
4	16.4	74.5	35311	1	US-08-629-600-1
5	16.4	74.5	35311	3	US-09-076-761-1
6	16.2	73.6	909	4	US-09-134-000C-3167
7	16.2	73.6	2358	4	US-09-134-000C-3285
8	15.8	71.8	660	4	US-09-489-039A-560
9	15.6	70.9	231	4	US-09-543-681A-2772
10	15.6	70.9	300	4	US-09-543-681A-1478
11	15.6	70.9	440	4	US-09-513-998C-35185
12	15.6	70.9	599	4	US-09-270-767-3552
13	15.6	70.9	599	4	US-09-270-767-18834
14	15.6	70.9	601	4	US-09-949-016-31352
15	15.6	70.9	601	4	US-09-949-016-69221
16	15.6	70.9	601	4	US-09-949-016-81370
17	15.6	70.9	601	4	US-09-949-016-127420
18	15.6	70.9	601	4	US-09-949-016-177125
19	15.6	70.9	601	4	US-09-949-016-177126
20	15.6	70.9	601	4	US-09-949-016-177127
21	15.6	70.9	678	3	US-09-134-001C-2551
22	15.6	70.9	826	3	US-09-221-0178-1008
23	15.6	70.9	850	3	US-08-617-860B-34
24	15.6	70.9	1068	4	US-09-543-681A-2972
25	15.6	70.9	1734	4	US-09-248-796A-4704
26	15.6	70.9	2109	4	US-09-248-796A-2999
27	15.6	70.9	2394	4	US-09-540-236-893

28	15.6	70.9	3069	3	US-08-335-865J-7	Sequence 7, Appl1
29	15.6	70.9	3805	3	US-09-513-729B-10	Sequence 10, Appl
30	15.6	70.9	3805	4	US-09-023-655-1443	Sequence 1443, Ap
31	15.6	70.9	4071	3	US-09-513-037C-5	Sequence 5, Appl1
32	15.6	70.9	4071	4	US-09-746-801A-5	Sequence 5, Appl1
33	15.6	70.9	4285	4	US-09-949-016-689	Sequence 23, Appl
34	15.6	70.9	4308	4	US-09-394-142B-23	Sequence 23, Appl
35	15.6	70.9	4643	2	US-08-605-106-6	Sequence 6, Appl1
36	15.6	70.9	5061	3	US-09-355-160D-1	Sequence 1, Appl1
37	15.6	70.9	5061	4	US-10-092-219-1	Sequence 1, Appl1
38	15.6	70.9	5296	4	US-09-949-016-2362	Sequence 2362, Ap
39	15.6	70.9	8302	3	US-09-234-827B-1	Sequence 1, Appl1
40	15.6	70.9	10204	4	US-09-949-016-14104	Sequence 14104, A
41	15.6	70.9	10482	3	US-09-322-478-23	Sequence 23, Appl
42	15.6	70.9	10482	4	US-09-586-106D-23	Sequence 23, Appl
43	15.6	70.9	14066	4	US-09-601-198-56	Sequence 56, Appl
44	15.6	70.9	26709	4	US-09-949-016-17520	Sequence 17520, A
45	15.6	70.9	34279	4	US-09-596-002-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1

US-09-313-294A-108  
; Sequence 108, Application US/09313294A  
; Patent No. 6476212  
; GENERAL INFORMATION:  
; APPLICANT: Ialugdi, Raghunath V.  
; APPLICANT: Ito, Laura Y.  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
; FILE REFERENCE: PL-0017 US  
; CURRENT APPLICATION NUMBER: US/09/313,294A  
; CURRENT FILING DATE: 1999-05-14  
; NUMBER OF SEQ ID NOS: 7600  
; SOFTWARE: PERL Program  
; SEQ ID NO 108  
; LENGTH: 133  
; TYPE: DNA  
; ORGANISM: Zea mays  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. 6476212 700548545H1  
; NAME/KEY: unsure  
; LOCATION: 8  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-313-294A-108

Query Match 78.2%; Score 17.2; DB 4; Length 133;  
Best Local Similarity 86.4%; Pred. No. 1.2e+02;  
Matches 19; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TCGTCGTTTTTCGTCGTTTTT 22  
|||||  
Db 72 TCGTCGTTGTCGTGTGTTT 93  
|||||

RESULT 2

US-09-107-532A-292/c  
; Sequence 292, Application US/09107532A  
; Patent No. 6583375  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 18:45:43 ; Search time 578 Seconds  
(without alignments)  
225.856 Million cell updates/sec

Title: US-10-613-228A-1

Perfect score: 22

Sequence: 1 tcgtcgcttttcgtcgctttt 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5401638 seqs, 2966923429 residues

Total number of hits satisfying chosen parameters: 10803276

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09C\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	22	100.0	22	US-10-613-228A-1
2	22	100.0	22	US-10-816-220-152
3	20.4	92.7	618	US-10-363-345A-34257
C 4	20.4	92.7	618	US-10-363-345A-34258
5	20.4	92.7	619	US-10-363-345A-40287
C 6	20.4	92.7	619	US-10-363-345A-40288
7	20.4	92.7	761	US-10-363-345A-2025
C 8	20.4	92.7	761	US-10-363-345A-2026
9	20.4	92.7	1024	US-10-363-345A-7203
C 10	20.4	92.7	1024	US-10-363-345A-7204
11	19.4	88.2	511	US-10-363-345A-28015

C 12	19.4	88.2	511	18	US-10-363-345A-28016	Sequence 28016, A
C 13	19.4	88.2	523	18	US-10-363-345A-35059	Sequence 35059, A
C 14	19.4	88.2	523	18	US-10-363-345A-35060	Sequence 35060, A
C 15	19.4	88.2	524	18	US-10-363-345A-40325	Sequence 40325, A
C 16	19.4	88.2	524	18	US-10-363-345A-40326	Sequence 40326, A
C 17	19.4	88.2	610	18	US-10-363-345A-23153	Sequence 23153, A
C 18	19.4	88.2	610	18	US-10-363-345A-23154	Sequence 23154, A
C 19	19.4	88.2	838	18	US-10-363-345A-32131	Sequence 32131, A
C 20	19.4	88.2	838	18	US-10-363-345A-32132	Sequence 32132, A
C 21	19.4	88.2	839	18	US-10-363-345A-1741	Sequence 1741, A
C 22	19.4	88.2	839	18	US-10-363-345A-1742	Sequence 1742, A
C 23	19.4	88.2	885	18	US-10-363-345A-18089	Sequence 18089, A
C 24	19.4	88.2	885	18	US-10-363-345A-18090	Sequence 18090, A
C 25	19.4	88.2	1267	18	US-10-363-345A-27261	Sequence 27261, A
C 26	19.4	88.2	1267	18	US-10-363-345A-27262	Sequence 27262, A
C 27	19.4	88.2	3673778	16	US-10-312-841-2	Sequence 2, Appli
C 28	19.4	86.4	920	18	US-10-363-345A-20285	Sequence 20285, A
C 29	19.4	86.4	920	18	US-10-363-345A-20286	Sequence 20286, A
C 30	18.8	85.5	523	18	US-10-363-345A-17495	Sequence 17495, A
C 31	18.8	85.5	523	18	US-10-363-345A-17496	Sequence 17496, A
C 32	18.8	85.5	524	18	US-10-363-345A-15599	Sequence 15599, A
C 33	18.8	85.5	524	18	US-10-363-345A-15600	Sequence 15600, A
C 34	18.8	85.5	525	18	US-10-363-345A-35507	Sequence 35507, A
C 35	18.8	85.5	525	18	US-10-363-345A-35508	Sequence 35508, A
C 36	18.8	85.5	553	18	US-10-363-345A-39041	Sequence 39041, A
C 37	18.8	85.5	553	18	US-10-363-345A-39042	Sequence 39042, A
C 38	18.8	85.5	561	18	US-10-363-345A-35971	Sequence 35971, A
C 39	18.8	85.5	561	18	US-10-363-345A-35972	Sequence 35972, A
C 40	18.8	85.5	651	18	US-10-363-345A-34913	Sequence 34913, A
C 41	18.8	85.5	651	18	US-10-363-345A-34914	Sequence 34914, A
C 42	18.8	85.5	712	18	US-10-363-345A-27451	Sequence 27451, A
C 43	18.8	85.5	712	18	US-10-363-345A-27452	Sequence 27452, A
C 44	18.8	85.5	759	18	US-10-363-345A-3263	Sequence 3263, A
C 45	18.8	85.5	759	18	US-10-363-345A-3264	Sequence 3264, A

ALIGNMENTS

RESULT 1

US-10-613-228A-1  
; Sequence 1, Application US/10613228A  
; Publication No. US20040092472A1  
; GENERAL INFORMATION:  
; APPLICANT: KRIGER, ARTHUR M  
; TITLE OF INVENTION: NUCLEIC ACID COMPOSITIONS FOR STIMULATING IMMUNE RESPONSES  
; FILE REFERENCE: C1037.70045US00  
; CURRENT APPLICATION NUMBER: US/10/613,228A  
; CURRENT FILING DATE: 2003-07-03  
; PRIOR APPLICATION NUMBER: US 60/394,193  
; PRIOR FILING DATE: 2002-07-03  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 22  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Oligodeoxynucleotide  
US-10-613-228A-1

Query Match 100.0%; Score 22; DB 17; Length 22;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCGTCGTTTTTCGTCGCGTTTTT 22

Db 1 TCGTCGTTTTTCGTCGCGTTTTT 22

RESULT 2

US-10-816-220-152  
; Sequence 152, Application US/10816220

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 17:17:45 ; Search time 3444 Seconds  
(without alignments)  
243.152 Million cell updates/sec

Title: US-10-613-228A-1

Perfect score: 22

Sequence: 1 tgcgtgttttttcgtcggtttt 22

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hrc:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gsl1:\*

9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20.4	92.7	433	4	BG428594 602494726
C 2	19.4	88.2	957	9	AG126335 Pan trogl
C 3	19.4	88.2	969	9	AG084147 Pan trogl
C 4	19.4	86.4	473	7	CN197708 TGESTZy15
C 5	18.8	85.5	698	1	AV895065 AV895065
C 6	18.8	85.5	750	9	CL653482 PRI0118C
C 7	18.8	85.5	755	5	BP006348 BP006348
C 8	18.8	85.5	793	9	CL659802 PRI0135b
C 9	18.8	85.5	1173	9	CG745933 P039-1-A0
C 10	18.8	85.5	1201	3	CR704325 Tetraodon
C 11	18.8	85.5	1372	9	CG751386 P045-4-D0
C 12	18.4	83.6	506	2	AV955361 AV955361
C 13	18.4	83.6	517	2	AV967757 AV967757
C 14	18.4	83.6	651	5	BW207212 BW207212
C 15	18.4	83.6	652	5	BW336085 BW336085
C 16	18.4	83.6	661	5	BW362385 BW362385
C 17	18.4	83.6	701	5	BP003979 BP003979
C 18	18.4	83.6	706	5	BP003979 BP003979
C 19	18.4	83.6	711	5	BM435032 BM435032
C 20	18.4	83.6	1101	9	CNS000Y0
C 21	18.4	83.6	236	7	W66251 TGESTZy2c0
C 22	17.8	80.9	246	6	CB365542 ZF001-P00
C 23	17.8	80.9	285	7	CN198693 TGESTZy16
C 24	17.8	80.9	342	6	CB353501 ZF001-P00

C 25	17.8	80.9	362	1	AV678562
C 26	17.8	80.9	374	1	AV679629
C 27	17.8	80.9	376	5	BW103452
C 28	17.8	80.9	378	1	AV894777
C 29	17.8	80.9	396	6	CB352916
C 30	17.8	80.9	444	5	BW589623
C 31	17.8	80.9	471	1	AV679587
C 32	17.8	80.9	498	5	BW575333
C 33	17.8	80.9	498	7	CN769410
C 34	17.8	80.9	499	1	AV887894
C 35	17.8	80.9	504	2	AV422790
C 36	17.8	80.9	512	4	BG892389
C 37	17.8	80.9	513	7	CN769717
C 38	17.8	80.9	525	1	AV885063
C 39	17.8	80.9	541	4	BG307575
C 40	17.8	80.9	544	5	BW104367
C 41	17.8	80.9	545	5	BW187581
C 42	17.8	80.9	555	5	BQ420126
C 43	17.8	80.9	559	6	CB353813
C 44	17.8	80.9	563	4	B1706183
C 45	17.8	80.9	564	2	BE200845

## ALIGNMENTS

RESULT 1  
BG428594/c  
LOCUS 602494726F1 NIH\_MGC\_75 Homo sapiens cDNA clone IMAGE:4608342 5',  
DEFINITION mRNA sequence.  
ACCESSION BG428594  
VERSION BG428594.1 GI:13335100  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 433)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsapbs-remail.nih.gov  
Tissue Procurement: CLONTECH Laboratories, Inc.  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1CM1350 row: d column: 07  
High quality sequence stop: 132.

FEATURES  
Location/Qualifiers  
1..433  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4608342"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_75"  
/note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattggcc); Site 2: SfiI (ggccattggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTGGC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.5-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

ORIGIN

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 14:14:33 ; Search time 3888 Seconds  
(without alignments)  
274.181 Million cell updates/sec

Title: US-10-613-228A-1  
Perfect score: 22  
Sequence: 1 tcgtcggttttcgtcggtttt 22

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_ats: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vi: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20.4	92.7	226153	2 AC073704	Mus muscu
2	19.4	88.2	349880	6 AX344573	Sequence
3	18.4	85.5	5826	6 AX356381	Sequence
4	18.8	85.5	6130	6 AX251400	Sequence
5	18.8	85.5	6130	6 AX345031	Sequence
6	18.8	85.5	7728	6 AX344979	Sequence
7	18.8	85.5	7728	6 AX348500	Sequence
8	18.8	85.5	8951	6 AX345697	Sequence
9	18.8	85.5	13133	6 AX344227	Sequence
10	18.4	83.6	20	6 AX104184	Sequence
11	18.4	83.6	20	6 AX355698	Sequence
12	18.4	83.6	20	6 AX547237	Sequence
13	18.4	83.6	5369	6 CQ000212	Sequence
14	18.4	83.6	146491	2 AC019950	Sequence
15	18.4	83.6	182601	3 AC009904	Sequence
16	18.4	83.6	236375	3 AE003708	Sequence
17	18.4	83.6	310967	1 AE016869	Sequence
18	18.4	81.8	149269	2 BX897667	Sequence
19	17.8	80.9	1527	5 BC056691	Sequence

20	17.8	80.9	2290	3	AB060934	Spirula s
21	17.8	80.9	2322	8	AY139685	Porphyra
22	17.8	80.9	2381	6	AX347239	Sequence
23	17.8	80.9	7195	6	AX277866	Sequence
24	17.8	80.9	7195	6	AX323551	Sequence
25	17.8	80.9	11622	6	AX345576	Sequence
26	17.8	80.9	34216	3	U57054	Sequence
27	17.8	80.9	99176	5	BX255893	Caenorhabdi
28	17.8	80.9	172585	2	CR847782	zebrafish
29	17.8	80.9	215917	2	AC006764	Danio rer
30	17.8	80.9	253250	1	AP005339	Caenorhab
31	17.8	80.9	300045	1	AE016803	Vibrio vu
32	17.8	80.9	303121	1	AE016766	Sequence
33	17.4	79.1	4110	6	AX598855	Sequence
34	17.4	79.1	5774	6	AX278031	Sequence
35	17.4	79.1	5774	6	AX323798	Sequence
36	17.4	79.1	5938	6	AX344811	Sequence
37	17.4	79.1	6067	6	AX344680	Sequence
38	17.4	79.1	6124	8	AB010408	Arabidops
39	17.4	79.1	7004	6	AX277956	Sequence
40	17.4	79.1	7004	6	AX323651	Sequence
41	17.4	79.1	7110	6	CQ006871	Sequence
42	17.4	79.1	7110	6	AX251243	Sequence
43	17.4	79.1	7110	6	AX251990	Sequence
44	17.4	79.1	7110	6	AX346458	Sequence
45	17.4	79.1	7110	6	AX349019	Sequence

## ALIGNMENTS

AC073704 226153 bp DNA linear HTG 29-JUN-2000  
Mus musculus clone RP23-175112, WORKING DRAFT SEQUENCE, 35  
unordered pieces.

AC073704  
AC073704.1 GI:8810321  
HTG; HTGS PHASE1; HTGS DRAFT.  
Mus musculus (house mouse)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

DOE Joint Genome Institute.  
1 (bases 1 to 226153)  
Sequencing of Mouse  
Unpublished  
2 (bases 1 to 226153)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (29-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: http://www.jgi.doe.gov

Project Information  
Center Project Name: 1804276  
Center clone name: RPCI-23\_175112

Summary Statistics  
Consensus Quality: 200795 bases at least Q40  
Consensus Quality: 211634 bases at least Q30  
Consensus Quality: 213977 bases at least Q20  
Estimated insert size: 205410; agarose-fp estimation  
Quality coverage: 7.81 in Q20 bases; agarose-fp estimation  
Quality coverage: 7.2 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 35 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 9, 2005, 14:06:53 ; Search time 535 Seconds  
(without alignments)  
243.429 Million cell updates/sec

Title: US-10-613-228A-1

Perfect score: 22

Sequence: 1 tcgtcgcttttcgtgcgttttt 22

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*

- 1: Geneseqn1980s:\*
- 2: Geneseqn1990s:\*
- 3: Geneseqn2000s:\*
- 4: Geneseqn2001as:\*
- 5: Geneseqn2001bs:\*
- 6: Geneseqn2002as:\*
- 7: Geneseqn2002bs:\*
- 8: Geneseqn2003as:\*
- 9: Geneseqn2003bs:\*
- 10: Geneseqn2003cs:\*
- 11: Geneseqn2003ds:\*
- 12: Geneseqn2004as:\*
- 13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	22	12 ADI16210	ADI16210 Immunosti
2	22	100.0	22	12 ADK19243	Adk19243 Immunosti
3	22	100.0	22	12 ABO44307	Ado44307 Nucleotid
4	20.4	92.7	618	6 ABO47667	Abq47667 Oligonucl
5	20.4	92.7	618	6 ABO47666	Abq47666 Oligonucl
6	20.4	92.7	619	6 ABO53697	Abq53697 Oligonucl
7	20.4	92.7	619	6 ABO53696	Abq53696 Oligonucl
8	20.4	92.7	761	6 ABO15435	Abq15435 Oligonucl
9	20.4	92.7	761	6 ABO15434	Abq15434 Oligonucl
10	20.4	92.7	1024	6 ABO20612	Abq20612 Oligonucl
11	20.4	92.7	1024	6 ABO20613	Abq20613 Oligonucl
12	19.4	88.2	511	6 ABO41425	Abq41425 Oligonucl
13	19.4	88.2	511	6 ABO41424	Abq41424 Oligonucl
14	19.4	88.2	523	6 ABO48469	Abq48469 Oligonucl
15	19.4	88.2	523	6 ABO48468	Abq48468 Oligonucl
16	19.4	88.2	524	6 ABO53734	Abq53734 Oligonucl
17	19.4	88.2	524	6 ABO53735	Abq53735 Oligonucl
18	19.4	88.2	610	6 ABO36562	Abq36562 Oligonucl
19	19.4	88.2	610	6 ABO36563	Abq36563 Oligonucl
20	19.4	88.2	838	6 ABO45540	Abq45540 Oligonucl

c 21	19.4	88.2	838	6	ABQ45541	Abq45541 Oligonucl
c 22	19.4	88.2	839	6	ABQ15151	Abq15151 Oligonucl
c 23	19.4	88.2	839	6	ABQ15150	Abq15150 Oligonucl
c 24	19.4	88.2	885	6	ABQ31498	Abq31498 Oligonucl
c 25	19.4	88.2	885	6	ABQ31499	Abq31499 Oligonucl
c 26	19.4	88.2	1267	6	ABQ40670	Abq40670 Oligonucl
c 27	19.4	88.2	1267	6	ABQ40671	Abq40671 Oligonucl
c 28	19	86.4	23	12	ADK19257	Adk19257 Immunosti
c 29	19	86.4	23	12	ADK19276	Adk19276 Immunosti
c 30	19	86.4	920	6	ABQ33695	Abq33695 Oligonucl
c 31	19	86.4	920	6	ABQ33694	Abq33694 Oligonucl
c 32	18.8	85.5	523	6	ABQ30904	Abq30904 Oligonucl
c 33	18.8	85.5	523	6	ABQ30905	Abq30905 Oligonucl
c 34	18.8	85.5	524	6	ABQ29009	Abq29009 Oligonucl
c 35	18.8	85.5	524	6	ABQ29008	Abq29008 Oligonucl
c 36	18.8	85.5	525	6	ABQ48916	Abq48916 Oligonucl
c 37	18.8	85.5	525	6	ABQ48917	Abq48917 Oligonucl
c 38	18.8	85.5	553	6	ABQ52450	Abq52450 Oligonucl
c 39	18.8	85.5	553	6	ABQ52451	Abq52451 Oligonucl
c 40	18.8	85.5	561	6	ABQ49381	Abq49381 Oligonucl
c 41	18.8	85.5	561	6	ABQ49380	Abq49380 Oligonucl
c 42	18.8	85.5	651	6	ABQ48323	Abq48323 Oligonucl
c 43	18.8	85.5	651	6	ABQ48322	Abq48322 Oligonucl
c 44	18.8	85.5	712	6	ABQ40860	Abq40860 Oligonucl
c 45	18.8	85.5	712	6	ABQ40861	Abq40861 Oligonucl

ALIGNMENTS

RESULT 1

ADI16210

ID ADI16210 standard; DNA; 22 BP.

XX AC ADI16210;

XX DT 22-APR-2004 (first entry)

XX DE Immunostimulatory oligodeoxynucleotide ODN 10106 SEQ ID NO:141.

XX ds; immunostimulatory; antibacterial; antiallergic; antiasthmatic;  
KW cytostatic; virucide; fungicide; antiparasitic; interleukin antagonist;  
KW gene therapy; infectious disease; allergy; asthma; cancer.  
XX OS Unidentified.  
XX PN WO2004005476-A2.  
XX PD 15-JAN-2004.  
XX PF 03-JUL-2003; 2003WO-US021113.  
XX PR 03-JUL-2002; 2002US-0393880P.  
XX PR 03-JUL-2002; 2002US-0394090P.  
XX PR 03-JUL-2002; 2002US-0394091P.  
XX PR 03-JUL-2002; 2002US-0394164P.  
XX PR 03-JUL-2002; 2002US-0394193P.  
XX (COLE-) COLEY PHARM GROUP INC.  
XX Krieg AM;  
XX WPI; 2004-091353/09.  
XX New immunostimulatory nucleic acid molecule composition comprising CpG  
PT motifs, useful for diagnosing, preventing and/or treating infectious  
XX diseases, allergies, asthma and cancers.  
XX Claim 1; SEQ ID NO 141; 257pp; English.  
XX The invention relates to a novel composition comprising an  
CC immunostimulatory nucleic acid molecule. A composition of the invention  
CC has antibacterial, antiallergic, antiasthmatic, cytostatic, virucide,